Claims

What is claimed is:

1. A method for managing cached data objects, the method comprising the steps of:

storing at least one data object in a central cache;

replicating the at least one data object in at least one local cache, the at least one local cache being in communication with the central cache;

maintaining, in accordance with the central cache, a directory describing content of the at least one local cache; and

in response to a change to cached data, the central cache updating at least a portion of content stored thereon and coordinating an update of at least a portion of content replicated on the at least one local cache using the directory to determine at least one data object in the at least one local cache which should be updated in accordance with the change.

2. The method of claim 1, further comprising the steps of:

a client requesting a data object from the at least one local cache; and

in response to the data object not being found on the at least one local cache, requesting the data object from the central cache.

3. The method of claim 2, further comprising the step of, in response to the data object not being found on the central cache, obtaining the data object from a remote server.

4. The method of claim 3, further comprising the step of, in response to obtaining the data object from the remote server, storing the data object in at least one of the central cache and the at least one local cache.

5. The method of claim 1, further comprising the step of applying a cache replacement policy to the at least one local cache in order to minimize utilization of a resource.

6. The method of claim 5, wherein the resource is the central cache.

7. The method of claim 6, wherein the cache replacement policy calculates a metric associated with a desirability of caching a data object, the metric being correlated with a product of a frequency with which the data object is accessed and a cost for the central cache to provide the object divided by the size of the data object.

8. The method of claim 5, wherein the at least one local cache and the central cache are connected by a network and the resource is the network.

9. The method of claim 8, wherein the cache replacement policy calculates a metric associated with a desirability of caching a data object, the metric being correlated with a frequency with which the data object is accessed.

10. The method of claim 5, further comprising the step of applying at least one other cache replacement policy to the at least one local cache in order to minimize utilization of at least one other resource.

11. The method of claim 10, further comprising the step of dynamically increasing a frequency with which a cache replacement policy designed to minimize utilization of a resource is applied in response to the resource becoming a bottleneck.

12. The method of claim 1, further comprising the step of applying a cache replacement policy to the at least one local cache in order to minimize a latency for obtaining a data object.

13. The method of claim 12, wherein the cache replacement policy calculates a metric associated with a desirability of caching a data object, the metric correlated with a frequency with which the data object is accessed divided by the data object size.

14. A method of determining a time for an event to occur, comprising the steps of:

determining a threshold time value t1;

determining if a frequency of an event occurring is low;

in response to the frequency being low, causing the event to occur; and

in response to the event not occurring after t1 has elapsed, causing the event to occur.

15. The method of claim 14, wherein the step of determining if a frequency of an event is low further comprises determining if a time between successive events one of equals and exceeds a second threshold time value t0.

16. The method of claim 14, further comprising the step of locking a data object by a process wherein the event comprises unlocking the data object and successive events comprise accesses to the data object by the process.

17. The method of claim 16, further comprising the step of increasing t1 in response to the process accessing the data object at a high rate.

18. The method of claim 16, further comprising the step of decreasing t1 in response to frequent requests to lock the data object.

19. The method of claim 16, wherein the data object is locked by a cache.

20. The method of claim 15, further comprising the steps of:

locking a data object by a process in which the event comprises unlocking the data object and the successive events comprise updates to the data object by the process; and

decreasing t0 in response to frequent requests to lock the data object.

21. A system for managing cached data objects, the system comprising:

a central cache for storing at least one data object and maintaining a directory describing content of at least one local cache; and

at least one local cache in communication with the central cache for replicating the at least one data object;

wherein, in response to a change to cached data, the central cache updates at least a portion of content stored thereon and coordinates an update of at least a portion of content replicated on the at least one local cache using the directory to determine at least one data object in the at least one local cache which should be updated in accordance with the change.

22. The system of claim 21, wherein a cache replacement policy is applied to the at least one local cache in order to minimize utilization of a resource.

23. The system of claim 21, wherein at least one other cache replacement policy is applied to the at least one local cache in order to minimize utilization of at least one other resource.

24. The system of claim 23, wherein a frequency, with which a cache replacement policy designed to minimize utilization of a resource is applied in response to the resource becoming a bottleneck, is dynamically increased.

25. The system of claim 21, wherein a cache replacement policy is applied to the at least one local cache in order to minimize a latency for obtaining a data object.

26. Apparatus for determining a time for an event to occur, the apparatus comprising:

at least one processor operative to: (i) determine a threshold time value t1; (ii) determine if a frequency of an event occurring is low; (iii) in response to the frequency being low, cause the event to occur; and (iv) in response to the event not occurring after t1 has elapsed, cause the event to occur; and

memory, operatively coupled to the at least one processor, for storing at least a portion of results of one or more operations performed by the at least one processor.

27. An article of manufacture for managing cached data objects, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

storing at least one data object in a central cache;

replicating the at least one data object in at least one local cache, the at least one local cache being in communication with the central cache;

maintaining, in accordance with the central cache, a directory describing content of the at least one local cache; and

in response to a change to cached data, the central cache updating at least a portion of content stored thereon and coordinating an update of at least a portion of content replicated on the at least one local cache using the directory to determine at least

one data object in the at least one local cache which should be updated in accordance with the change.

28. An article of manufacture for determining a time for an event to occur, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

determining a threshold time value t1;

determining if a frequency of an event occurring is low;

in response to the frequency being low, causing the event to occur; and

in response to the event not occurring after t1 has elapsed, causing the event to occur.